THE USE OF SINGLE TRYPTOPHAN VARIANTS TO STUDY PROTEIN FOLDING AND STABILITY

A Senior Honors Thesis

by

JENNIFER NATALIE DULIN

Submitted to the Office of Honors Programs
& Academic Scholarships
Texas A&M University
in partial fulfillment of the requirements of the

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOWS

April 2004

Major: Biochemistry

THE USE OF SINGLE TRYPTOPHAN VARIANTS TO STUDY PROTEIN FOLDING AND STABILITY

A Senior Honors Thesis

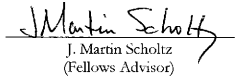
by

JENNIFER NATALIE DULIN

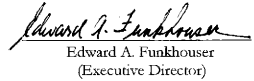
Submitted to the Office of Honors Programs
& Academic Scholarships
Texas A&M University
in partial fulfillment for the designation of

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOW

Approved as to style and content by:



J. Martin Scholtz
(Fellows Advisor)



Edward A. Funkhouser
(Executive Director)

April 2004

Major: Biochemistry

ABSTRACT

The Use of Single Tryptophan Variants to Study Protein Folding and Stability. (April 2004)

Jennifer Natalie Dulin
Department of Biochemistry/Biophysics
Texas A&M University

Fellows Advisor: Dr. J. Martin Scholtz
Department of Medical Biochemistry and Genetics

Studies on the kinetics of protein folding of the histidine-containing phosphocarrier protein (HPr) from the thermophile *Bacillus stearothermophilus* (*Bst*) will contribute much to the understanding of the origins of its enhanced thermal stability. Although circular dichroism (CD) is available to us as a method by which to study the unfolding properties of HPr, the more effective method of fluorescence spectroscopy improves on the high signal-to-noise ratio of CD. Tryptophan is the best amino acid to use as a probe by which to monitor unfolding by fluorescence; however, wild type *Bst*HPr contains no tryptophan residues. In order to utilize this technique, three tryptophan-containing variants of HPr (F6W, F29W, and Y37W) were designed and expressed in hopes of finding a variant protein which possesses similar properties to wild type HPr so that it may be used as a model. The usefulness of these probes were analyzed by fluorescence emission scans, and it was determined that each of the three residues acted as good probes by which to monitor the folding and unfolding of HPr by fluorescence. Subsequently, experiments were performed where urea denaturation was monitored by circular dichroism and fluorescence. ΔG and *m*-values for each variant protein were obtained by analysis of the resultant urea denaturation curves. Results indicated that Y37W has the closest ΔG of stabilization and *m*-value to the

wild type protein, suggesting that it can be used as a model protein for kinetic studies of the folding mechanisms of HPr.

ACKNOWLEDGMENTS

I thank Dr. Marty Scholtz for allowing me, as an inexperienced freshman, to join the ranks of his lab and learn the research process firsthand. His guidance and patience are much appreciated, and his professionalism has influenced me to become a better researcher. I also thank Abbas Razvi for his boundless patience, his instruction in the laboratory, and his friendship. I thank all of the members of the Scholtz and Pace labs for providing me with a learning environment of assistance and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Proteins play an exceedingly significant role in living organisms; they control vital cell processes such as metabolism, membrane traffic, and transport, and they constitute much of the structure of cells. *In vivo*, every protein exists in a unique three-dimensional folded structure that is necessary for proper function; function is often diminished or lost entirely if the structure is abnormal. An in-depth understanding of the relatedness of structure and function is important for applications such as the design of drugs for medical purposes and enzymes for industrial applications. Furthermore, modern genomics has sequenced the genomes of many different organisms, unearthing a wealth of new genetic information. This sequence information is only useful when function is attributed.

Structure-determination techniques such as nuclear magnetic resonance (NMR) and X-ray crystallography are widely used for this purpose, but are often quite time- and labor-intensive. The goal of determining a protein's three-dimensional folded structure from examination of its primary amino acid sequence is one that, if achieved, could provide a superior alternative method to the classical approaches to structure determination. However, a protein's native state is stabilized by several factors including hydrogen bonding, electrostatic interactions, and hydrophobic interactions. Consequently it is very difficult to predict the exact three-dimensional (native) structure of a protein given only its primary sequence. Therefore, an understanding of the interactions involved in protein folding is key in order to translate cryptic gene sequences into useful information.

This thesis follows the style and format of *Protein Science*.

Histidine-containing phosphocarrier protein (HPr) is involved in the uptake of sugars in bacteria. It is a small globular protein with no disulfide bonds or prosthetic groups, so it serves as an ideal model protein for research in the area of protein folding. Studies have been performed on HPr from organisms such as *Escherichia coli* and *Bacillus subtilis*, yielding much information about the kinetics of folding and the thermal stability of these proteins. Recently, our lab has begun studies on HPr from *Bacillus stearothermophilus* (*Bst*), a thermophile that lives at 60°C and thus possesses greater thermal protein stability than that of mesophilic organisms. This research will analyze the thermodynamics of *Bst*HPr, ultimately contributing to the determination of the origins of its enhanced stability as compared with HPr from the mesophile *B. subtilis*.

To determine the conformational stability of *Bst*HPr, two techniques will be used to analyze the stability as a function of urea concentration. Circular dichroism (CD) provides a way of detecting the change in ellipticity at 222 nm of our protein as urea disrupts the native state by solvating the protein's backbone. From this data, a curve of CD signal as a function of urea concentration can be produced, and fitting the curve to an appropriate function allows us to determine free energy (ΔG) of stabilization for the protein. Fluorescence spectroscopy is another method of studying stability. The aromatic amino acid residues (phenylalanine, tyrosine, and tryptophan) report on their local environment; as a protein unfolds and the normally buried aromatic residues become more solvated, fluorescence properties change. Therefore, the degree of foldedness of a protein can be determined by measurement of the fluorescence properties of these residues.

In this technique, the fluorescence of aromatic amino acids is measured as a function of increasing urea concentration, also providing valuable information about the

unfolding of a protein. For the purposes of fluorescence, tryptophan has superior properties than phenylalanine and tyrosine, as it shows significant signal change upon protein denaturation, making it an ideal probe for these studies. Thus, it is desirable that a protein to be studied by this method contains a well-placed tryptophan probe.

Of the several techniques employed to study the kinetics of protein folding and unfolding, the best method available to us is fluorescence spectroscopy. This method is much more effective because it has a much better signal-to-noise ratio than CD; however, a tryptophan probe must be present within the protein in order to achieve optimum fluorescence data. Therefore, the first objective of this study is to create tryptophan variants of our model HPr by site-directed mutagenesis. The sites of tryptophan substitution have been selected by inspection of the three-dimensional structure of the protein by molecular modeling. Figure 1 shows a model of HPr with tryptophan residues introduced at positions 6, 29, and 37. Once these protein variants were successfully expressed and purified, their unfolding kinetics were characterized by CD and fluorescence. This allows us to probe the local environment of each tryptophan residue and to determine the effects on protein folding and stability caused by the mutation. Ultimately, in creating these variants, the goal is to find one that has the most similar properties to those of *Bst*HPr, so studies with that variant can be used to provide insight into the folding mechanisms of the wild type protein. Two important properties that were tested for are that the tryptophan in the variant protein serves as a good fluorescence probe, and that the variant behaves like *Bst*HPr.

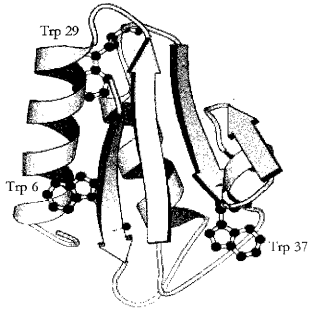


Fig. 1. *Bs*HPr with labeled residues 6, 29, and 37 mutated to tryptophan and displayed as ball and stick.

MATERIALS AND METHODS

Design and creation of the variants

Oligonucleotide primers containing the desired mutations in the HPr gene were designed and ordered from Integrated DNA Technologies. To insert these mutations into the plasmid encoding wild type *Bst*HPr, site-directed mutagenesis was performed using the QuikChange Site-Directed Mutagenesis Kit produced by Stratagene, by running a PCR reaction on a Perkin Elmer GeneAmp PCR System 2400 thermocycler. The primers used are as follows: F6W, 5'- GAA AAA ACG TGG AAA GTC GTT TC-3'; F29W, 5'- ACG GCG AGC AAA TGG AAC AGC GAA ATC-3'; Y37W, 5'- CCA GCT TGA GTG GAA CGG CAA-3'. Parental DNA was then digested, and the presence of plasmid DNA was verified by gel electrophoresis.

The DNA obtained from PCR amplification was transformed according to QuikChange protocol into *E. coli* XL1-Blue supercompetent cells and plated out on LB agar plates containing 50 μg/mL ampicillin in order to select for cells containing the *Bst* vector, which carries an ampicillin-resistance marker. After 16 hours of incubation at 37°C, one well-isolated colony was selected from each plate, inoculated into sterile LB media with ampicillin, and allowed to grow overnight with shaking at 37°C. Plasmid DNA was extracted from these cells using the QIAprep Spin Miniprep Kit produced by Qiagen. After extraction, sequences were confirmed at the Gene Technologies Laboratory (Department of Biology, Texas A&M University), ensuring that the desired mutations were indeed present in the *Bst* plasmid DNA.

Protein expression and purification

The variant HPr proteins were expressed in *E. coli* according by a modification of the methods of Reizer et al. (1989). The mutant plasmid DNA was transformed into *E. coli* ES7R *trp, rpsL, recA, tet* electrocompetent cells by electroporation, plated out on LB agar plates containing ampicillin and tetracycline, and incubated at 37°C for 16 hours. 5-mL ampicillin-containing LB broth cultures were grown from isolated colonies obtained from the LB plates, and incubated at 37°C overnight. 6-L LB broth preps were then prepared and inoculated with the cell culture; ampicillin and tetracycline were added at final concentrations of 50 mg/L and 5 mg/L, respectively, and the flasks incubated at 37°C with shaking overnight.

Cells were harvested by centrifugation and resuspended in 20 mM Tris, 2 mM EDTA. Cell suspension was passed through the french press at 1100-1200 psi in order to lyse the cells; the lysate was centrifuged and the supernatant was diluted in 20 mM Tris, 2mM EDTA buffer. Undesired proteins and other particles were precipitated out of solution with 22.6% $(NH_4)_2SO_4$ by weight and the suspension was then centrifuged. The supernatant was dialyzed against 4 L of a 20 mM Tris, 2 mM EDTA buffer overnight with one buffer change, and the dialyzed solution was applied to a DEAE ion-exchange resin column which had been equilibrated with an HPr buffer at pH 8.4. Flow-through protein was lyophilized, resuspended in dH_2O, and applied to a Sephadex G-50 gel filtration column with a 20 mM NH_4HCO_3 buffer. Fractions of the column eluent were collected and analyzed by SDS-PAGE. Fractions containing HPr were pooled and lyophilized. Protein purity was confirmed by SDS-PAGE and mass spectrometry.

Urea denaturations and determination of stability

CD at 222 nm was employed to monitor the urea-induced unfolding of each tryptophan variant, using an Aviv 62DS as well as an Aviv 202SF spectropolarimeter. Urea stock solutions were prepared for each run with sodium phosphate buffer (at a final concentration of 10 mM) at pH 7.0, the concentrations of which were determined by refractive index measurements (Pace, 1986). 10-mL urea stock solutions with a final protein concentration of 6.6 μM were prepared and the concentrations of these samples were also determined by refractive index. Similarly, the cuvettes contained 2 mL of protein solution and 10 mM sodium phosphate buffer at the same protein concentration. The instrument made urea additions in serial increments, and after each addition and equilibration, CD signal was recorded. It is possible to fit the curve for the denaturation data by using the equation:

$$y = \frac{\{(y_f + m_f[D]) + (y_u + m_u[D]) \cdot \exp[m([D] - [D]_{1/2})/RT\}}{(1 + \exp[m([D] - [D]_{1/2})/RT)}, \quad (1)$$

where m measures the steepness of the transition, $[D]_{1/2}$ measures the midpoint of the transition, and the pre- and post-transition regions can also be characterized by two parameters y_f and y_u for the intercepts, and m_f and m_u for the slopes of the folded and unfolded protein, respectively (Pace & Scholtz, 1990). In this manner, a urea-induced denaturation curve was constructed. This curve can be used to find the ΔG (free energy of stabilization) and m-value (slope of the dependence of ΔG on denaturant concentration) of the transition.

Urea-induced denaturations were also monitored by fluorescence with λ_{ex} of 280 nm using an Aviv 202SF spectropolarimeter with a cutoff filter of 350 nm for emission. 10-mL

urea stock solutions with final concentrations of 1.5 μM were prepared at pH 7.0, and their concentrations were determined by refractive index. The cuvettes contained protein at the same final concentration in 10 mM sodium phosphate buffer. Small aliquots of the urea-protein solution were then added directly by the instrument to the cuvette with stirring at a constant temperature. Sequential additions of urea were made and fluorescence of the tryptophan residues at 280 nm was monitored at 25°C.

RESULTS

Fluorescence emission scans

Fluorescence emission scans were collected from 300 – 400 nm on proteins *Bst*F6W, F29W, and Y37W HPr with an excitation wavelength of 280 nm. The buffer-corrected spectra for each variant are shown in Figure 2. These scans were performed in order to determine whether the tryptophan residues were in well placed positions to serve as good fluorescence probes. These spectra show that each variant protein emits drastically different fluorescence signal in the denatured state than that of the native state, indicating that the introduced tryptophan residues will serve as good probes of foldedness of theirs respective host proteins. This is beneficial for our purposes, as each of these residues, placed in different regions of the protein, will act as good probes to follow protein folding and unfolding by fluorescence.

Stability measurements from urea denaturations

The conformational stability of the variant proteins at pH 7.0 in 10 mM sodium phosphate has been determined from urea denaturation curves at various temperatures, by both CD and fluorescence. CD signal at 222 nm and fluorescence emission at 280 nm were monitored as a function of urea concentration, and this data was converted into denaturation curves by constructing a best-fit line based on Equation 1. Selected curves for each variant attained by both methods are shown in Figures 3 through 8. All denaturation curves obtained had sufficient data points in the pre- and post-transition regions, and showed well-defined cooperative transitions from the folded to the denatured state. These

qualities made for the typical sigmoidal curve shape which can be observed in Figures 3 through 8. All experiments were repeated for reproducibility. The quality of the curve fits to the data was ascertained by the relative proximity of R-values (see tables these figures) to unity.

These curves were analyzed with the linear extrapolation method (LEM) as described by Pace (1986) and Santoro and Bolen (1992), yielding ΔG values and *m*-values (the linear relationship between ΔG and denaturant concentration). *M*-values can be derived from the curve data using the equation:

$$m = R \cdot m5 \cdot T, \tag{2}$$

where *m*-value is expressed in cal mol^{-1} M^{-1} , m5 is the slope of the transition, *R* is the gas constant (1.987 cal mol^{-1} K^{-1}), and temperature is expressed in degrees Kelvin. Once *m*-value is determined, ΔG of stabilization is derived by the equation:

$$\Delta G = \frac{m\text{-value} \cdot m6}{1000}, \tag{3}$$

where ΔG is expressed in kcal mol^{-1} and m6 represents the midpoint of the transition. Table 1 illustrates the average ΔG values and *m*-values for each variant as well as wild type *Bst*HPr, determined by fitting the data from the urea denaturation curves to Equations 2 and 3.

Selection of the ideal tryptophan variant

For each protein variant, ΔG and *m*-values were analyzed for their closeness to those of the wild type protein, as well as for the agreement in data from CD and fluorescence. As shown in Table 1, ΔG of stabilization determined by CD for wild type *Bst*HPr is 8.23 kcal mol^{-1}, and *m*-value is 1000 cal mol^{-1} M^{-1}. The protein with the closest agreement in these

parameters to wild type HPr is Y37W, with only a –0.05 kcal mol^{-1} change from wild type ΔG and the least deviation in *m*-value as measured by both CD and fluorescence. The ΔG values determined for F29W were 7.45 by fluorescence and 7.37 by CD; moreover, the ΔG determined for F6W were the most deviant from wild type at 6.63 by fluorescence and 7.50 by CD, respectively. Standard deviations calculated from these measurements were 0.07 for F6W, 0.18 for F29W, and 0.08 for Y37W. Furthermore, the accuracy of these measurements is verified by the excellent agreement between the ΔG values of the protein determined independently by each method. Although standard deviation for free energy for F6W were the smallest, the ΔG obtained by CD for F6W is 0.87 kcal mol^{-1} greater than the value obtained by fluorescence. Similarly, the difference in ΔG values for F29 W is 0.08 kcal mol^{-1}. This data indicates possible departure from two-state folding mechanisms; making F6W and F29W less than ideal candidates for studies in stability. This data demonstrates that *Bst*Y37W HPr has the most similar properties to those of wild type *Bst*HPr, suggesting that it can be used as a model protein for kinetic studies of the folding mechanisms of HPr.

Table 1. Analysis of urea-denaturation curves for wt BstHPr and Trp variants

Protein	UDC by Fluorescence		UDC by CD		Difference	$\Delta G_{WT} - \Delta G_{VAR}$ (kcal mol^{-1})	Standard deviation of ΔG
	avg ΔG (kcal mol^{-1})	m-value (cal mol^{-1} M^{-1})	avg ΔG (kcal mol^{-1})	avg m-value (cal mol^{-1} M^{-1})	$\Delta G_{FL} - \Delta G_{CD}$ (kcal mol^{-1})		
wt *Bst*	-	-	8.23	1000	-	0.00	-
*Bst*F6W	6.63	1023	7.50	1079	-0.87	0.73	0.07
*Bst*F29W	7.45	1055	7.37	1058	0.08	0.86	0.18
*Bst*Y37W	8.18	996	8.18	1032	0.00	0.05	0.08

[a]Experiments done on the Aviv 202SF with λ_{ex} of 280nm and a 350nm cutoff filter for emission. All experiments at 25°C and pH 7.0 in 10 mM sodium phosphate.

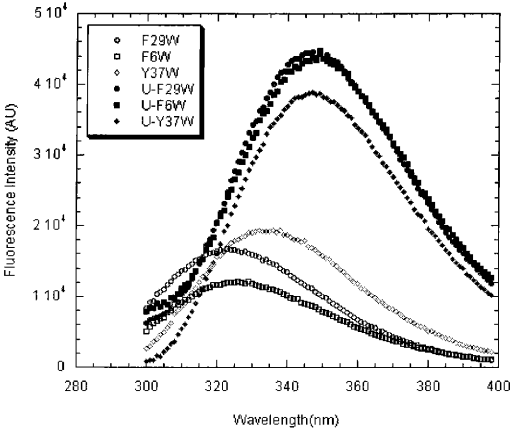


Fig. 2. Fluorescence emission scans for *Bst*F6W, F29W, and Y37W HPr. Open data points designate folded protein and closed data points designate unfolded protein. Protein denatured in 10.3 M urea, 10 mM sodium phosphate buffer, pH 7.0. Fluorescence signal was detected with a λ_{ex} of 280 nm.

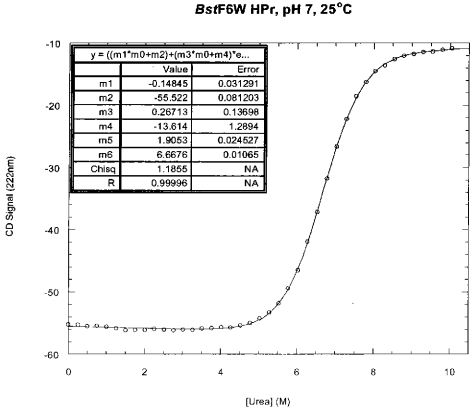


Fig. 3. Urea induced unfolding of *Bst*F6W HPr in 10 mM sodium phosphate buffer, pH 7.0 at 25°C measured by CD absorbance at 222 nm. The solid line is fit to the data.

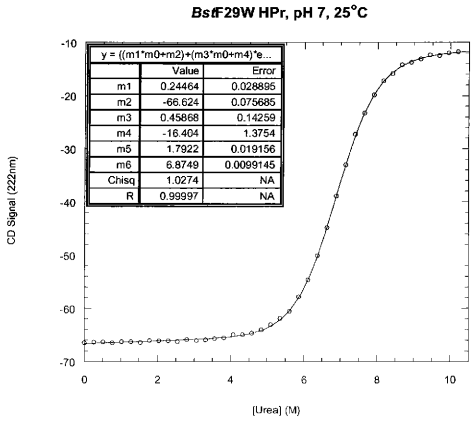


Fig. 4. Urea induced unfolding of *Bst*F29W HPr in 10 mM sodium phosphate buffer, pH 7.0 at 25°C measured by CD absorbance at 222 nm. The solid line is fit to the data.

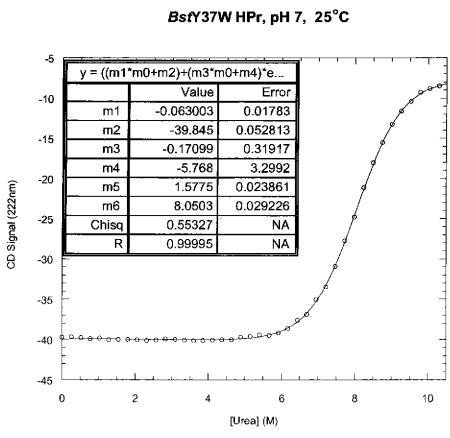


Fig. 5. Urea induced unfolding of *Bst*Y37W HPr in 10 mM sodium phosphate buffer, pH 7.0 at 25°C measured by CD absorbance at 222 nm. The solid line is fit to the data.

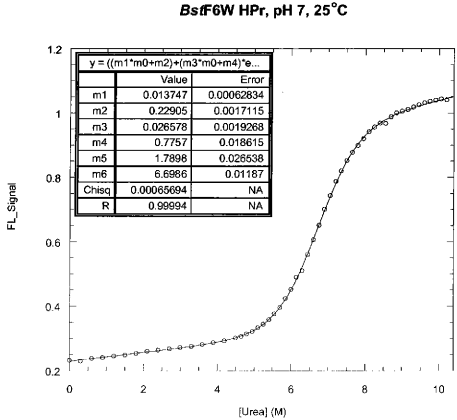


Fig. 6. Urea induced unfolding of *Bst*F6W HPr in 10 mM sodium phosphate buffer, pH 7.0. Fluorescence signal was detected using a 350 nm cut-off filter. The λ_{ex} was 280 nm.

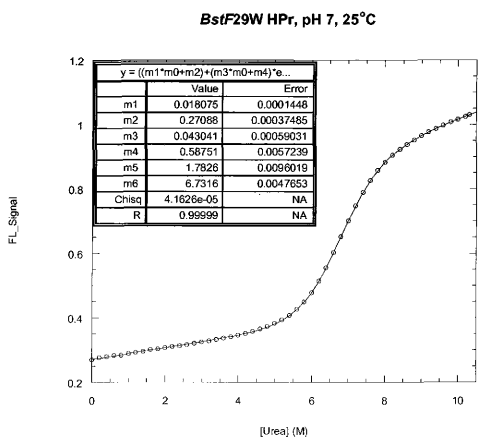


Fig. 7. Urea induced unfolding of *Bst*F29W HPr in 10 mM sodium phosphate buffer, pH 7.0. Fluorescence signal was detected using a 350 nm cut-off filter. The λ_{ex} was 280 nm.

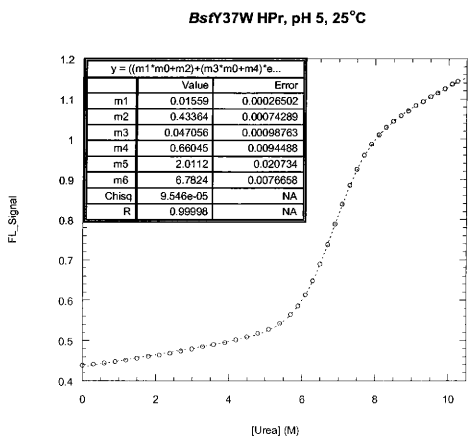


Fig. 8. Urea induced unfolding of *Bsf*Y37W HPr in 10 mM sodium phosphate buffer, pH 5.0. Fluorescence signal was detected using a 350 nm cut-off filter. The λ_{ex} was 280 nm.

SUMMARY

Kinetic studies on HPr from *Bacillus stearothermophilus* will help explain the enhanced thermal stability of the protein. In order to utilize fluorescence spectroscopy as a method by which to study the kinetics of HPr, tryptophan-containing variants of this protein were designed and assayed for two properties: (1) The quality of tryptophan residues as a fluorescence probe, and (2) The stability of the wild type protein must be preserved despite the introduction of the tryptophan residue; in other words, the residue must not disrupt the overall protein structure or afford any loss of stability upon the protein.

Fluorescence emission scans showed that *Bst*F6W, F29W, and Y37W HPr all serve as good probes because of the significant change in fluorescence signal from the native and denatured states of each protein. Conformational stability of the three variants was then tested by urea denaturation monitored by CD and fluorescence. Denaturation curves were produced by fitting Equation 1 to the data of CD or fluorescence signal versus urea concentration. Using Equations 2 and 3, these curves then were analyzed and the ΔG and *m*-values for each protein variant were obtained; the average values for the proteins obtained from several experiments are listed in Table 1.

As determined experimentally by these spectroscopic measurements, *Bst*Y37W HPr has the closest ΔG and *m*-value to wild type HPr. The deviation in free energy of stabilization from wild type protein is –0.05 kcal mol^{-1}, and the deviation in *m*-values determined by fluorescence and CD are –4 cal mol^{-1} M^{-1} and +32 cal mol^{-1} M^{-1}, respectively. Furthermore, there is an excellent agreement between ΔG obtained by fluorescence and CD

data. This data indicates that Y37W is a good fluorescent model for examining the kinetics of HPr.

Fluorescence emission scans can also be performed over several different concentrations of urea to provide information about the folding mechanism of a protein. Presence or absence of an isofluoric point in these data has been established as a criterion for a two state nature of the protein's folding mechanism. Future work with *Bs*Y37W HPr will be to perform fluorescence emission scans as a function of increasing urea concentrations in order to ascertain the two state nature of the protein, as per the methods of Azuaga et al. (2003).

REFERENCES

Azuaga AI, Canet D, Smeenk G, Berends R, Titgemeijer F, Duurkens R, Mateo PL, Scheek RM, Robillard GT, Dobson CM, van Nuland NA. 2003. Characterization of single-tryptophan mutants of histidine-containing phosphocarrier protein: evidence for local rearrangements during folding from high concentrations of denaturant. *Biochemistry 42*:4883-4895.

Pace CN. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol 131*:266-280.

Pace CN, Scholtz JM. 1990. Measuring the conformational stability of a protein. In: Creighton TE, Ed. *Protein structure: a practical approach*. 299-320.

Reizer J, Sutrina SL, Saier MH Jr, Stewart GC, Peterkofsky A, Reddy P. 1989. Mechanistic and physiological consequences of HPr(ser) phosphorylation on the activities of the phosphoenolpyruvate:sugar phosphotransferase system in gram-positive bacteria: Studies with site-specific mutants of HPr. *EMBO J 8*:2111-2120.

Santoro MM, Bolen DW. 1992. A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range. *Biochemistry 31*:4901-4907.

VITA

Jennifer Natalie Dulin

2221 N Belmont

Richmond, TX 77469

I will graduate from Texas A&M University in the spring of 2005 with a Bachelor's degree in Biochemistry and a minor in English. After graduation, I plan to attend graduate school to earn a Ph.D. in Biochemistry. Following graduation, I plan to pursue a career in research in the field of structural biology.

During my time at Texas A&M, I have received a Lechner Fellowship and a Director's Excellence Award. I have been named Distinguished Student of the TAMU College of Life Sciences. I will graduate with University and Foundation Honors, as well as the designations of University Undergraduate Research Fellow and Distinguished Biochemistry Research Scholar. I am also a member of Phi Eta Sigma National Honor Society. I served as an ASPIRE peer mentor from 2002 – 2003, and an Honors Invitational Program leader in 2002. Furthermore, I have participated in and organized community service projects such as 30 Hour Famine, Adopt-a-Family, and The Big Event.

From February 2002, I have worked as an undergraduate researcher in the lab of Dr. Marty Scholtz in the Texas A&M Department of Medical Biochemistry and Genetics. I plan to continue research in this lab until graduation.